

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

DJ
#2 30902
Priority Paper

Applicant(s): Tsuei-Chi YEH, et al.)
Serial No.: 09/884,287)
Filed: June 19, 2001) Our Ref: B-4218 618886-4)
For: "METHOD FOR IDENTIFYING)
AUTHORIZED USERS USING A)
SPECTROGRAM AND APPARATUS OF)
THE SAME") Date: December 10, 2001

CLAIM TO PRIORITY UNDER 35 U.S.C. 119

Commissioner of Patents and Trademarks
Box New Patent Application
Washington, D.C. 20231

RECEIVED

FEB 05 2002

Technology Center 2600

Sir:

[X] Applicants hereby make a right of priority claim under 35
U.S.C. 119 for the benefit of the filing date(s) of the
following corresponding foreign application(s):

<u>COUNTRY</u>	<u>FILING DATE</u>	<u>SERIAL NUMBER</u>
Taiwan, R.O.C.	27 December 2000	89128026

[] A certified copy of each of the above-noted patent
applications was filed with the Parent Application
No. _____.

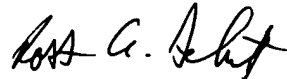
[X] To support applicants' claim, a certified copy of the above-
identified foreign patent application is enclosed herewith.

[] The priority document will be forwarded to the Patent Office
when required or prior to issuance.

I hereby certify that this correspondence
is being deposited with the United States
Postal Service with sufficient postage as
first-class mail in an envelope addressed
to the "Commissioner of Patents and
Trademarks, Washington, D.C. 20231",
on December 10, 2001 by Cheryl Liang.

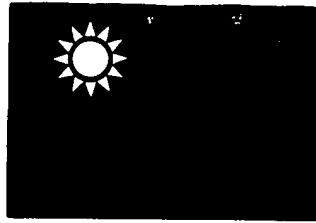


Respectfully submitted,



Ross A. Schmitt
Attorney for Applicant
Reg. No. 42,529

LADAS & PARRY
5670 Wilshire Boulevard
Suite 2100
Los Angeles, CA 90036
Telephone: (323) 934-2300
Telefax: (323) 934-0202



中華民國經濟部智慧財產局

INTELLECTUAL PROPERTY OFFICE
MINISTRY OF ECONOMIC AFFAIRS
REPUBLIC OF CHINA



茲證明所附文件，係本局存檔中原申請案的副本，正確無訛，
其申請資料如下：

This is to certify that annexed is a true copy from the records of this
office of the application as originally filed which is identified hereunder:

申請 日：西元 2000 年 12 月 27 日
Application Date

申請 案 號：089128026
Application No.

申請 人：華邦電子股份有限公司
Applicant(s)

RECEIVED

FEB 05 2002

Technology Center 2600

局 長

Director General

陳 明 邦

發文日期：西元 2001 年 8 月 13 日
Issue Date

發文字號：09011011920
Serial No.



申請日期：

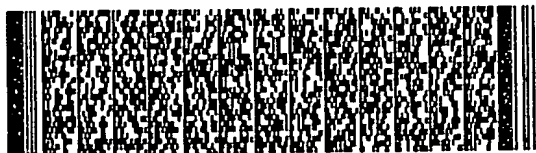
案號：

類別：

(以上各欄由本局填註)

發明專利說明書

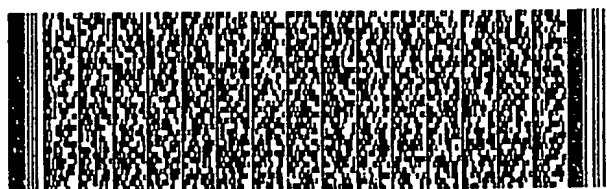
一、 發明名稱	中文	利用聲譜資訊辨識授權使用者的方法與其裝置
	英文	
二、 發明人	姓名 (中文)	1. 葉垂奇 2. 陳文源
	姓名 (英文)	1. 2.
	國籍	1. 中華民國 2. 中華民國
	住、居所	1. 新竹市科學里22鄰竹村二路4之2號 2. 新竹市東區仙宮里寶山路452巷5弄5號
三、 申請人	姓名 (名稱) (中文)	1. 華邦電子股份有限公司
	姓名 (名稱) (英文)	1.
	國籍	1. 中華民國
	住、居所 (事務所)	1. 新竹科學工業園區研新三路四號
	代表人 姓名 (中文)	1. 焦佑鈞
	代表人 姓名 (英文)	1.



四、中文發明摘要 (發明之名稱：利用聲譜資訊辨識授權使用者的方法與其裝置)

一種利用聲譜資訊辨識授權使用者的方法與其裝置，其利用不同使用者特有的聲譜資訊來辨識使用者的身份，以決定使用者是否經過授權。上述方法係包括下列步驟：(i)使用者發出語音後，偵測語音之終點；(ii)自上述語音的聲譜中取出語音的特徵；(iii)決定是否需要訓練，若是則將上述語音特徵當做一參考樣本，同時並設定一界限，否則進行下一步驟；(iv)將上述語音特徵與參考樣本進行圖樣比對；(v)根據上述比對結果，計算其間的距離；(vi)將上述計算的結果與設定的界限比較；(vii)依據上述比較的結果，決定上述使用者是否為一授權使用者。

英文發明摘要 (發明之名稱：)



本案已向

國(地區)申請專利

申請日期

案號

主張優先權

無

有關微生物已寄存於

寄存日期

寄存號碼

無

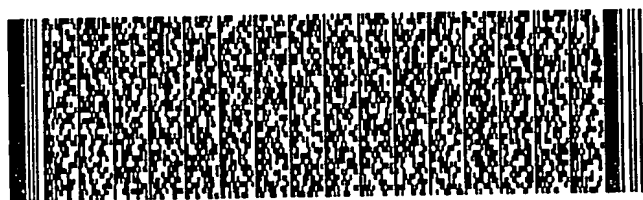
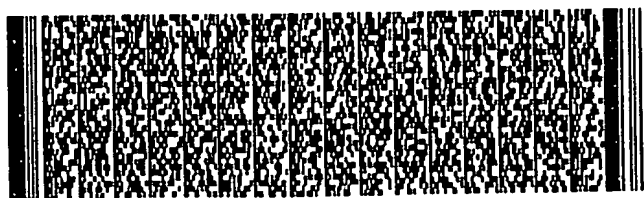
五、發明說明 (1)

本發明係有關於一種語音辨識方法及裝置，特別係有關於一種利用聲譜資訊辨識授權使用者的方法與其裝置。

隨著通信科技的發展，目前行動電話的使用有日益普遍的趨勢，行動電話確實使人們日常生活中的通信聯絡變得相當地便利。然而，行動電話的安全性卻始終存在著問題，未經授權的使用者可能在未經同意的情況下盜打行動電話，造成電話所有人的損失。

為了防止行動電話被盜打，一般行動電話手機上通常提供有密碼辨識的功能。也就是，在手機開機時，會先要求使用者輸入密碼，在密碼正確的情況下，使用者才能使用行動電話。可是此種方式要求使用者必須記住密碼，假使使用者忘記密碼或輸入了錯誤的密碼，則可能造成行動電話被鎖住而無法使用。再者，未經授權的使用者也有可能經由不同的途徑知悉密碼的內容，如此一來，此種密碼辨識的方式就未必能夠滿足行動電話安全上的需求。

因此，除了上述的方式外，習知方法中也有利用語音辨識來辨別說話者的技術，例如，在美國專利第 5,913,196 號中，其利用至少兩個聲音鑑定演算法(voice authentication algorithms)來分析說話者的聲音。另外，在美國專利第 5,499,288 號中，其主要是從說話者的聲音中取出時域特徵(heuristically-developed time domain features)及頻域資訊例如快速傅立葉轉換(FFT)係數，以做為主要特徵，再依據此主要特徵，分別依序找出第二及第三特徵。然後，利用這些特徵進行語音辨識的



五、發明說明 (2)

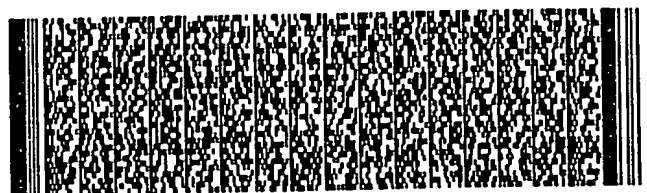
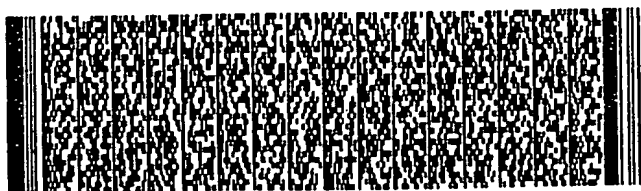
流程。至於，美國專利第5,365,574號，其與上述美國專利第5,499,288號類似，不過其另外提供了選擇性可調整的訊號界限。在美國專利第5,216,720號中，則是使用了LPC(linear predictive coding)分析來取得語音特徵，並使用DTW(dynamic time warping)法來對輸入語音特徵及參考語音特徵間的距離評分。上述各種習知技術雖均是引用語音辨識的方式來識別說話者，然其使用的方法均各不相同。當語音辨識的方式應用至行動電話上時，由於其要避免複雜龐大的硬體架構，所以上述習知方法在實施上往往有其困難。

有鑑於此，為了克服習知技術的缺點，本發明之目的即在於提出一種新的辨識授權使用者的方法與其裝置，其利用不同使用者特有的聲譜資訊來辨識使用者的身份，以決定使用者是否經過授權。

本發明之裝置具有簡單的架構，可滿足行動電話輕、薄、短、小的要求。

由於每個人說話的方式以及發音的器官，包括發音區域的結構、鼻腔的大小及聲帶的特徵等，均存在有天生的差異，使得每個人說話都包含了其獨特的資訊，因此本發明主要是利用聲譜分析方法將這些獨特的資訊自語音中取出，藉以辨識使用者。

本發明主要是在利用每一時框(frame)的主要值與使用者設定的界限進行比較，決定語音之始點與終點後，再利用Princen-Bradley的濾波器來轉換偵測到的語音訊號



五、發明說明 (3)

，以取得其對應的聲譜圖案。然後經由將得到的聲譜圖案與事先儲存的使用者之聲譜的參考樣本進行比對，以決定是否為授權的使用者。

圖式之簡單說明

第1圖係繪示依據本發明辨識電話之授權使用者之方法的流程圖。

第2圖係繪示依據本發明用以偵測語音之裝置的方塊圖。

第3圖係繪示依據本發明之預先加強處理之示意圖。

第4圖係繪示依據本發明決定短時間主要值之流程圖。

第5圖係繪示依據本發明偵測終點之步驟的流程圖。

第6圖係繪示依據本發明自聲譜中取出語音特徵之方法的流程圖。

第7圖係繪示本發明之利用聲譜資訊辨識授權使用者的裝置之方塊圖。

參考標號之說明

10~低通濾波器；

20~類比/數位轉換器；

30~數位訊號處理器；

40~記憶裝置。

實施例之說明

在本實施例中，以行動電話的使用者為例。請參閱第1圖，依據本發明辨識電話之授權使用者之方法係包括下



五、發明說明 (4)

列步驟：(i)步驟100，使用者發出語音後，偵測語音之終點；(ii)步驟110，自上述語音的聲譜中取出語音的特徵；(iii)步驟120，決定是否需要訓練，若是則進行至步驟122，將上述語音特徵當做一參考樣本，同時進行至步驟124，設定一界限，否則進行下一步驟；(iv)步驟130，將上述語音特徵與參考樣本進行圖樣比對；(v)步驟140，根據上述比對結果，計算其間的距離；(vi)步驟150，將上述計算的結果與設定的界限比較；(vii)步驟160，依據上述比較的結果，決定上述使用者是否為一授權使用者。

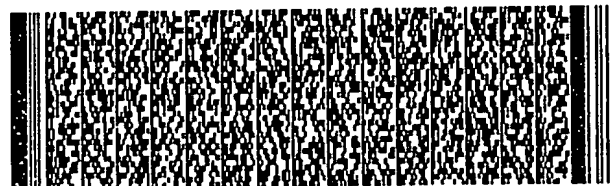
接下來，分別進一步說明上述各步驟的實施方式。

請參閱第2圖，上述偵測語音終點的方式係包括下列步驟：(i)步驟200，首先，由麥克風輸入的語音先經過一低通濾波器；(ii)步驟210，接著經過一類比/數位轉換器，對每一數位化的樣本，以8KHz的速率取樣，其解析度為8位元；(iii)步驟220，為了良好地取得語音的低振幅及高頻部分，數位化的資料通過一前端增強器(pre-emphasizer)；(iv)步驟230，取得主要值(majority magnitude)；(v)步驟240，比較每一時框(frame)的主要值與一預設的界限，以決定語音的始點與終點。

上述步驟200中，低通濾波器的頻率限制為3500Hz。

由於在本實施例中，前端增強的因子a選定為31/32，如此，一簡單的前端增強過程可由下列的運算完成：

$$y(n) = x(n) - ax(n-1) = x(n) - (31/32)x(n-1) = x(n) - x(n-1) + x(n-1)/32$$



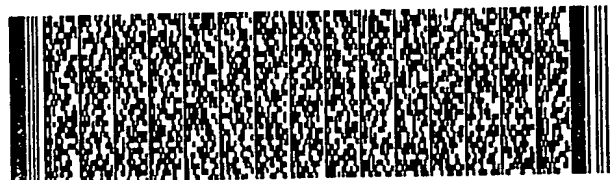
五、發明說明 (5)

因此在上述步驟220中，對數位化的資料進行前端增強的運算過程即如第3圖所示。

接著，前端增強的語音資料以時框(frame)為單位被分割，每一時框包括有160個樣本(0.02秒)。同時，取得一參數，即上述步驟230中之主要值，以描述振幅的特性。請參閱第4圖，上述取得主要值的過程係包括下列步驟：(i)步驟400，清除陣列 $ary[0], \dots, ary[127]$ ；(ii)步驟410，判斷語音資料 $y(n)$ 是否屬於目前的時框，若是則進行下一步驟，否則進行步驟430；(iii)步驟420，更新 $y(n)$ 的陣列值 $ary[|y(n)|]$ ，使 $ary[|y(n)|] = ary[|y(n)|] + 1$ ；(iv)步驟422，繼續下一個語音資料，使 $n = n + 1$ ，然後回到步驟410；(v)步驟430，求得每一語音資料之陣列值 $ary[0], \dots, ary[127]$ 的最大值的平均值 k ；(vi)步驟440，定義第 i 個時框的主要值 $mmag(i) = k$ ；(vii)步驟450，是否進行下一個時框，若是則進行步驟452，否則停止運算；(viii)步驟452，進行下一個時框的運算，使參數 $i = i + 1$ ，然後回到步驟400。

在上述取得主要值的過程中，對於每一個時框，計算每一個振幅位準的絕對值的總數。振幅位準中的大多數(great majority)被定義為目前的時框的主要值。

請參閱第5圖，上述步驟240中之決定語音始點與終點的流程係包括下列步驟：(i)步驟500，設定界限為20；(ii)步驟510，判斷是否偵測到始點，若是則進行步驟540，否則進行下一步驟；(iii)步驟520，判斷是否有連

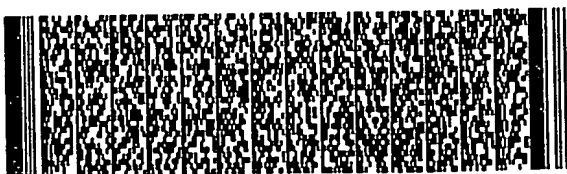


五、發明說明 (6)

續的三個主要值 $mmg(i-2)$, $mmg(i-1)$, $mmg(i)$ 均大於界限，若是則進行步驟530，否則進行下一步驟；(iv)步驟522，更新界限；(v)步驟524，使 $i = i+1$ ，然後回到步驟510；(vi)步驟530，已偵測到始點；(vii)步驟532，始點位於第 $i-2$ 個時框；(viii)步驟534，使 $k = 0$ ，然後進行到步驟524；(ix)步驟540，使 $k = k+1$ ；(x)步驟550，判斷 k 是否大於10，若是則進行下一步驟，否則回到步驟540；(xi)步驟560，判斷是否有連續的三個主要值 $mmg(i-2)$, $mmg(i-1)$, $mmg(i)$ 均小於界限，若是則進行步驟570，否則進行下一步驟；(xii)步驟562，使 $i = i+1$ ，然後回到步驟560；(xiii)步驟570，已偵測到終點；(vix)步驟580，終點位於第 $i-2$ 個時框，然後停止運算。

在上述終點偵測的流程中，一開始先設定背景噪音的界限為20。對每一個輸入的時框，計算其主要值，然後將其與預設的界限比較，以決定其是否為語音的一部分。假如連續三個時框的主要值均大於界限，表示已偵測到語音的始點；否則，目前的時框即被視為是新的背景噪音，且對界限進行更新。界限的更新程序可由下列方程式的運算完成：

$$\begin{aligned} \text{new_threshold} &= (\text{old_threshold} \times 31 + \text{new_input}) \\ &\div 32 \\ &= (\text{old_threshold} \times 32 - \text{old_threshold} \\ &+ \text{new_input}) \div 32 \\ &= \text{old_threshold} + (\text{new_input} - \end{aligned}$$

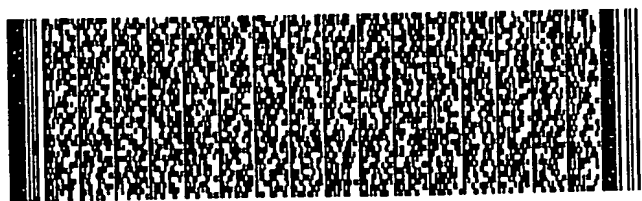


五、發明說明 (7)

$\text{old_threshold}) \div 32$

上述的除法運算可利用數位資料的移位運算加以完成。另外，由於假設對於一個聲音而言，至少會有0.3秒的時間。所以，對語音終點的偵測，在偵測到始點後10個時框才開始。又假如連續三個時框的主要值均小於界限，表示已偵測到語音的終點。

為了從聲譜中取得語音特徵，在本實施例中主要是利用Princen-Bradley的濾波器來轉換偵測到的語音訊號，以取得其對應的聲譜。有關Princen-Bradley的濾波器的說明可參閱John P. Princen及Alan Bernard Bradley, "Analysis/Synthesis Filter Bank Design Based On Time Domain Aliasing Cancellation," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, No. 5, Oct. 1986, pp. 1153-1161。請參閱第6圖，上述從聲譜中取得語音特徵的流程係包括下列步驟：(i)步驟600，首先定義時框長度 $K = 256$ ，時框速率 $M = 128$ ；(ii)步驟610，偵測到的聲音有 T 個PCM樣本 $x(n)$ ， $n = 0, \dots, T-1$ ；(iii)步驟620，利用Princen-Bradley過濾器 $X(k, m)$ 來計算聲譜，其中， $k = 0, \dots, K/2$ ， $m = 0, \dots, T/M$ ；(iv)步驟630，將 T/M 個向量平均區段化成 Q 個區段，並將第 q 個區段的向量平均而得到一新的向量 $Z(q) = Z(0, q), \dots, Z(K/2, q)$ ；(v)步驟640，搜尋區域的峰值，若 $Z(k, q) > Z(k+1, q)$ 且 $Z(k, q) > Z(k-1, q)$ ，則 $Z(k, q)$ 為一區域的峰值，設定 $W(k, q) = 1$ ，否則設定 $W(k, q) = 0$ ，



五、發明說明 (8)

其中， $k = 0, \dots, K/2$ ， $q = 0, \dots, Q-1$ ， W 是最後的特徵向量，然後停止運算。

在上述從聲譜中取得語音特徵的流程中，主要是利用 Princen-Bradley 的濾波器來轉換偵測到的語音訊號，以取得其對應的聲譜。假設一個時框有 K 個 PCM 樣本，且目前的時框有 M 個 PCM 樣本與下一個時框重疊。在本實施例中， K 及 M 分別被設為 256 與 128。如此，在第 m 個時框的第 k 個頻帶的訊號可利用下式加以計算：

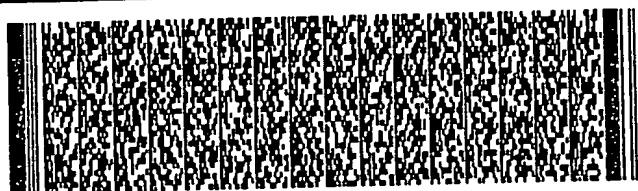
$$Y(k, m) = S y(n)h(mM-n+K-1)\cos(m\pi/2-2\pi(n+n_0)/K)$$

上述函數 h 中的係數可在前述 Princen 與 Bradley 的論文第九個表格中查到。 $Y(m) = Y(0, m), \dots, Y(K/2, m)$ 係涵蓋了 0Hz~4000Hz 的頻率範圍。假如被偵測到的語音具有 T 個 PCM 樣本， $Y(m)$ 的 $L(L=T/M)$ 個向量將被計算以得出 T 個 PCM 樣本的聲譜。 L 個向量會被平均地區段化成 Q 個區段。第 q 個區段的向量也被平均而得到一新的向量 $Z(q) =$

$Z(0, q), \dots, Z(K/2, q)$ 。接著，執行一區域峰值的搜尋子程式，經由設定 $W(k, q) = 1$ 以代表峰值，其他則設定 $W(k, q) = 0$ ，而標示出區域的峰值。最後，可得到一具有 $Q(K/2+1)$ 個位元的圖案以表示被偵測到之語音的聲譜。

最後，進行圖案的匹配及距離的運算。在參考樣本 RW (由 $RW(0), \dots, RW(Q)$ 形成) 及測試樣本 TW (由 $TW(0), \dots, TW(Q)$ 形成) 間的距離評分 (distance scoring) 可利用下式加以計算：

$$\text{dis} = S |TW(i, j) - RW(i, j)|, \text{ 其中, } i = 0, \dots, K/2,$$



五、發明說明 (9)

$j = 0, \dots, Q$

因為 $TW(i, j)$ 與 $RW(i, j)$ 的值不是1就是0，上式可以簡單地經由位元的運算而完成。在第1圖中的界限是由授權的使用者預先設定。假如上式所得到的dis不超過界限，則本發明之裝置即輸出一接受的指令。

請參閱第7圖，本發明之利用聲譜資訊辨識授權使用者的裝置係包括：一低通濾波器10，一類比/數位轉換器20，一數位訊號處理器30，及一記憶裝置40。

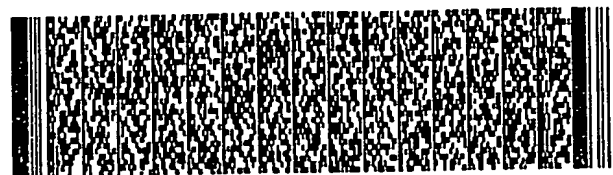
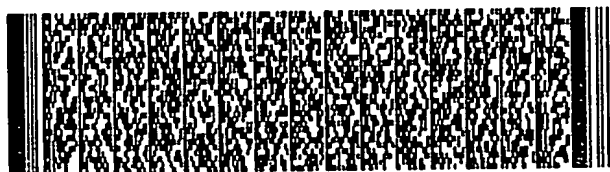
其中，上述低通濾波器10係用以限制輸入語音之頻率範圍。

上述類比/數位轉換器20係用以將輸入語音之類比訊號轉換成數位訊號，以便進行後續的處理。

上述數位訊號處理器30係接收經上述類比/數位轉換器20輸出的數位訊號，並進行前述第1圖中各步驟之運算。

上述記憶裝置40則係用以儲存界限與參考樣本等資料，藉以提供上述數位訊號處理器30運算之所需。

本發明雖以較佳實施例揭露如上，然其並非用以限定本發明，任何熟習此項技藝者，在不脫離本發明之精神和範圍內，仍可對語音偵測、聲譜特徵的取得、參考樣本的訓練、計算參考樣本與測試樣本間距離、及決定是否為授權使用者的方式等做些許的更動與潤飾，因此本發明之保護範圍當視後附之申請專利範圍所界定者為準。



六、申請專利範圍

1. 一種利用聲譜資訊辨識授權使用者的方法，包括下列步驟：

- (i) 使用者發出語音後，偵測語音之終點；
- (ii) 自上述語音的聲譜中取出語音的特徵；
- (iii) 決定是否需要訓練，若是則將上述語音特徵當做一參考樣本，同時並設定一界限，否則進行下一步驟；
- (iv) 將上述語音特徵與參考樣本進行圖樣比對；
- (v) 根據上述比對結果，計算其間的距離；
- (vi) 將上述計算的結果與設定的界限比較；
- (vii) 依據上述比較的結果，決定上述使用者是否為一授權使用者。

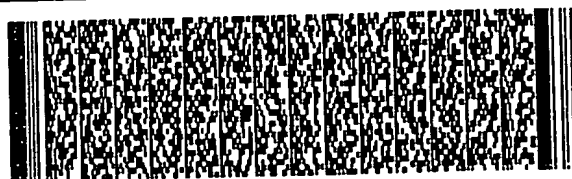
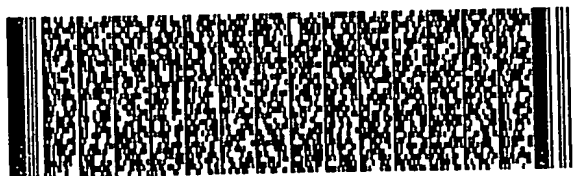
2. 如申請專利範圍第1項之方法，其中，上述步驟(i)中語音終點的偵測係包括下列步驟：

- (i) 由麥克風輸入的語音先經過一低通濾波器；
- (ii) 經過一類比/數位轉換器；
- (iii) 使數位化的資料通過一前端增強器；
- (iv) 取得主要值；

(v) 比較每一時框的主要值與一預設的界限，以決定語音的始點與終點。

3. 如申請專利範圍第1項之方法，上述取出語音特徵的方式係利用Princen-Bradley的濾波器來轉換偵測到的語音訊號，以取得其對應的聲譜。

4. 如申請專利範圍第2項之方法，其中，上述主要值係對每一時框計算其每一個振幅位準的絕對值的總數，並



六、申請專利範圍

且定義振幅位準中的最大多數為目前的時框的主要值。

5. 如申請專利範圍第2項之方法，其中，上述步驟(v)中決定語音之始點與終點的流程係包括下列步驟：

(i) 設定一界限；

(ii) 決定是否開始偵測起點，若否則進行下一步驟，否則進行至步驟(iv)；

(iii) 決定是否連續三個主要值均大於界限，若否則修正界限，並繼續量測下一主要值，且回到上述步驟(ii)，否則即表示已偵測到起點，繼續量測下一主要值，並回到步驟(ii)；

(iv) 延遲一段時間；

(v) 決定是否有連續三個主要值均小於界限，若否則繼續量測下一主要值，並重新回到步驟(v)，否則即表示已偵測到終點。

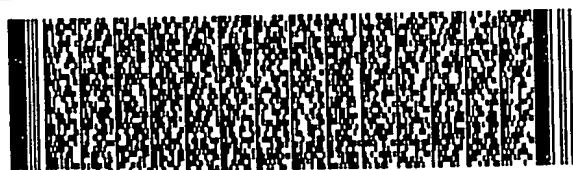
6. 一種利用聲譜資訊辨識授權使用者的裝置，包括：

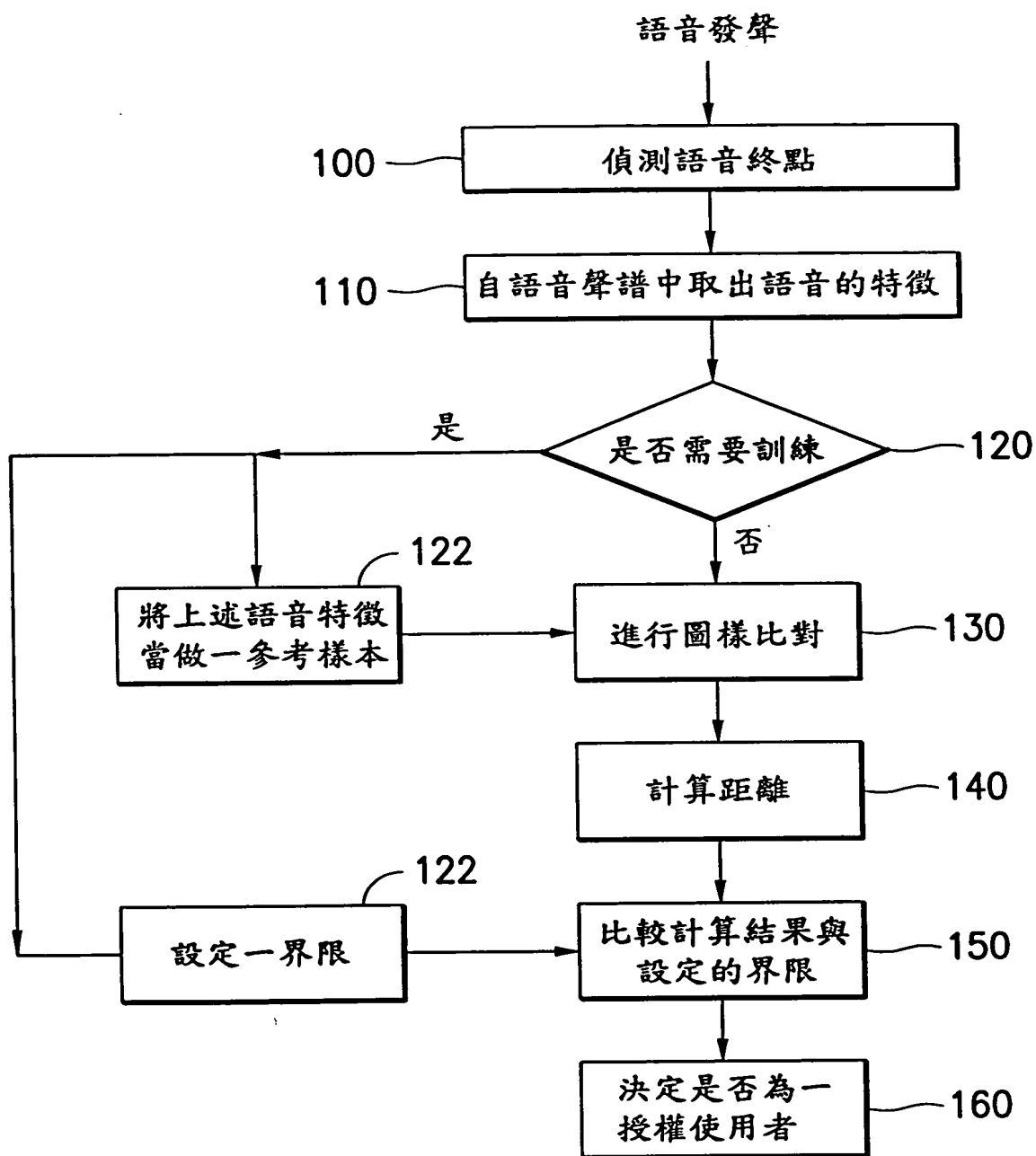
一低通濾波器，用以限制輸入語音之頻率範圍；

一類比/數位轉換器，用以將輸入語音之類比訊號轉換成數位訊號，以便進行後續的處理；

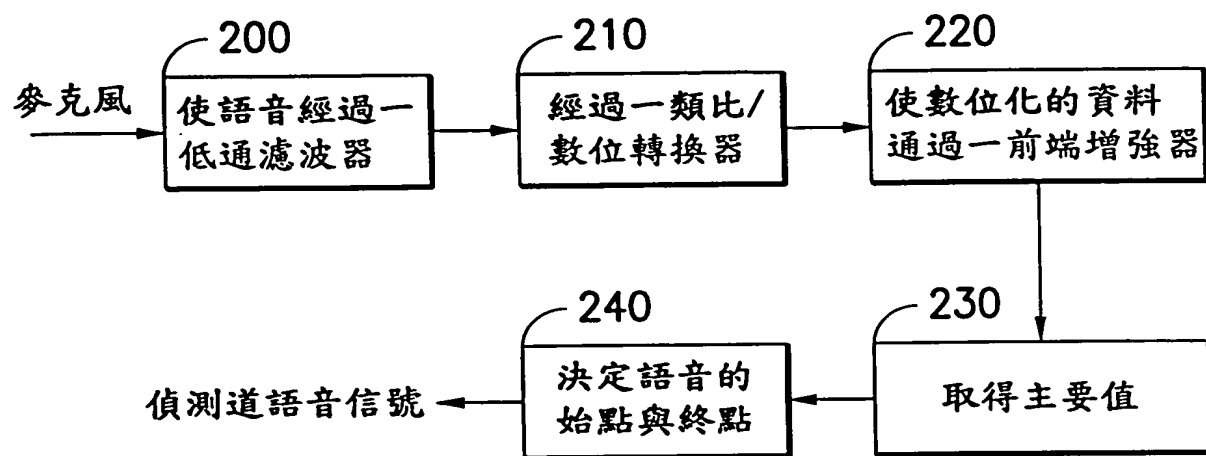
一數位訊號處理器，接收經上述類比/數位轉換器輸出的數位訊號，並進行申請專利範圍第1項之方法中各步驟之運算；及

一記憶裝置，用以儲存界限與參考樣本等資料，藉以提供上述數位訊號處理器運算之所需。

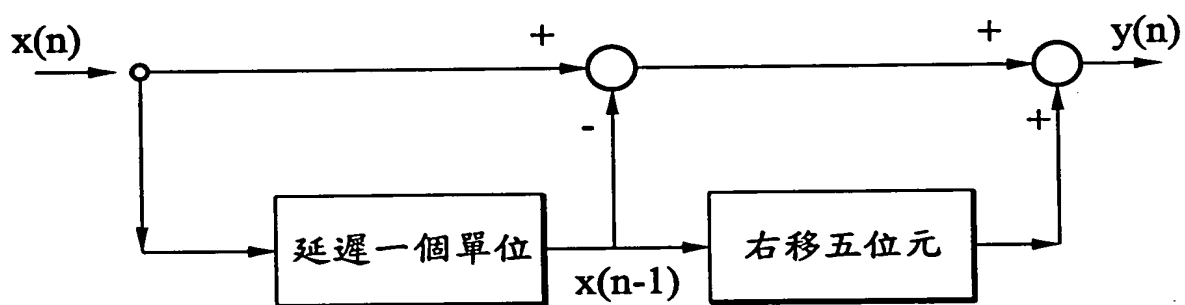




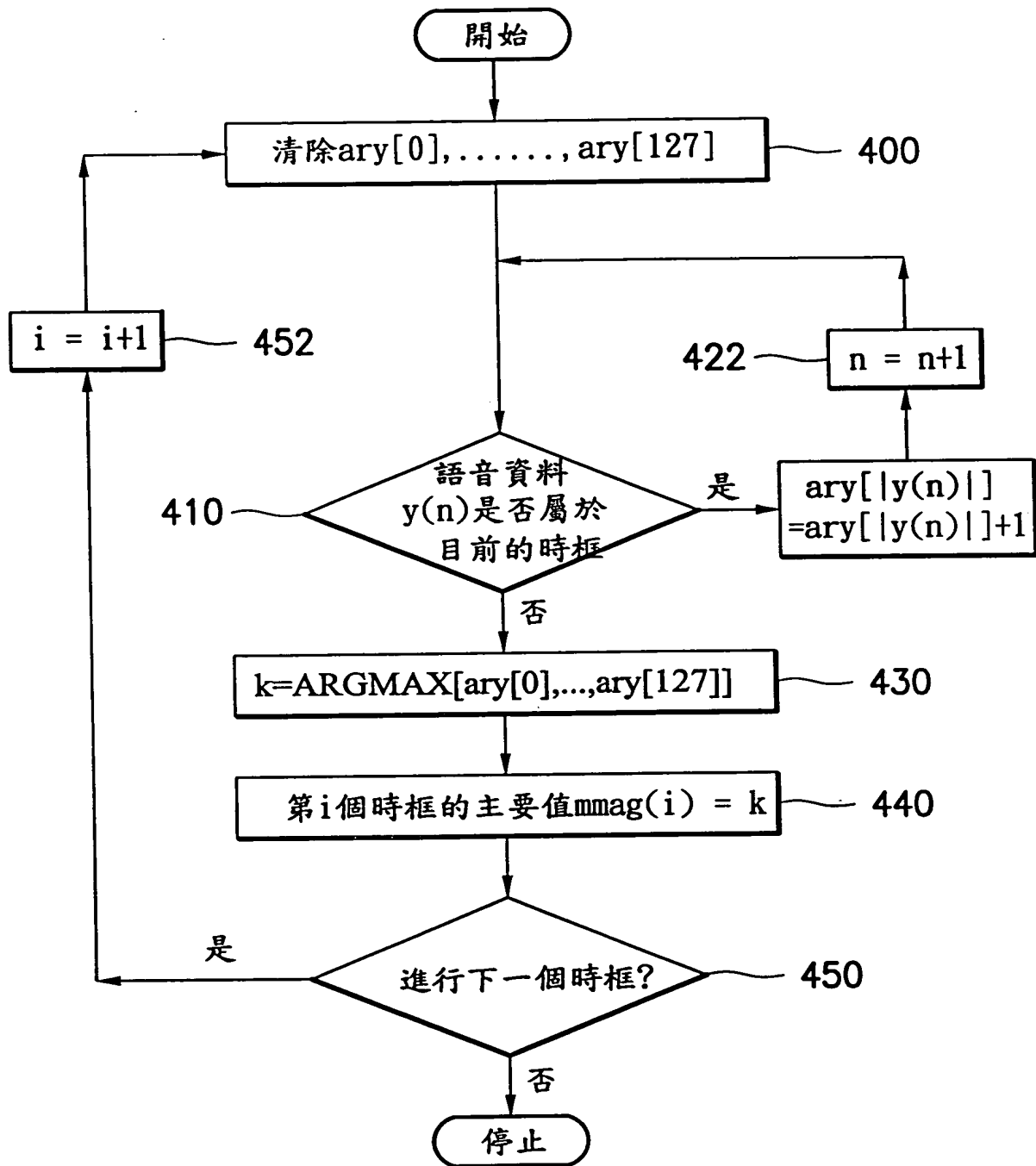
第 1 圖



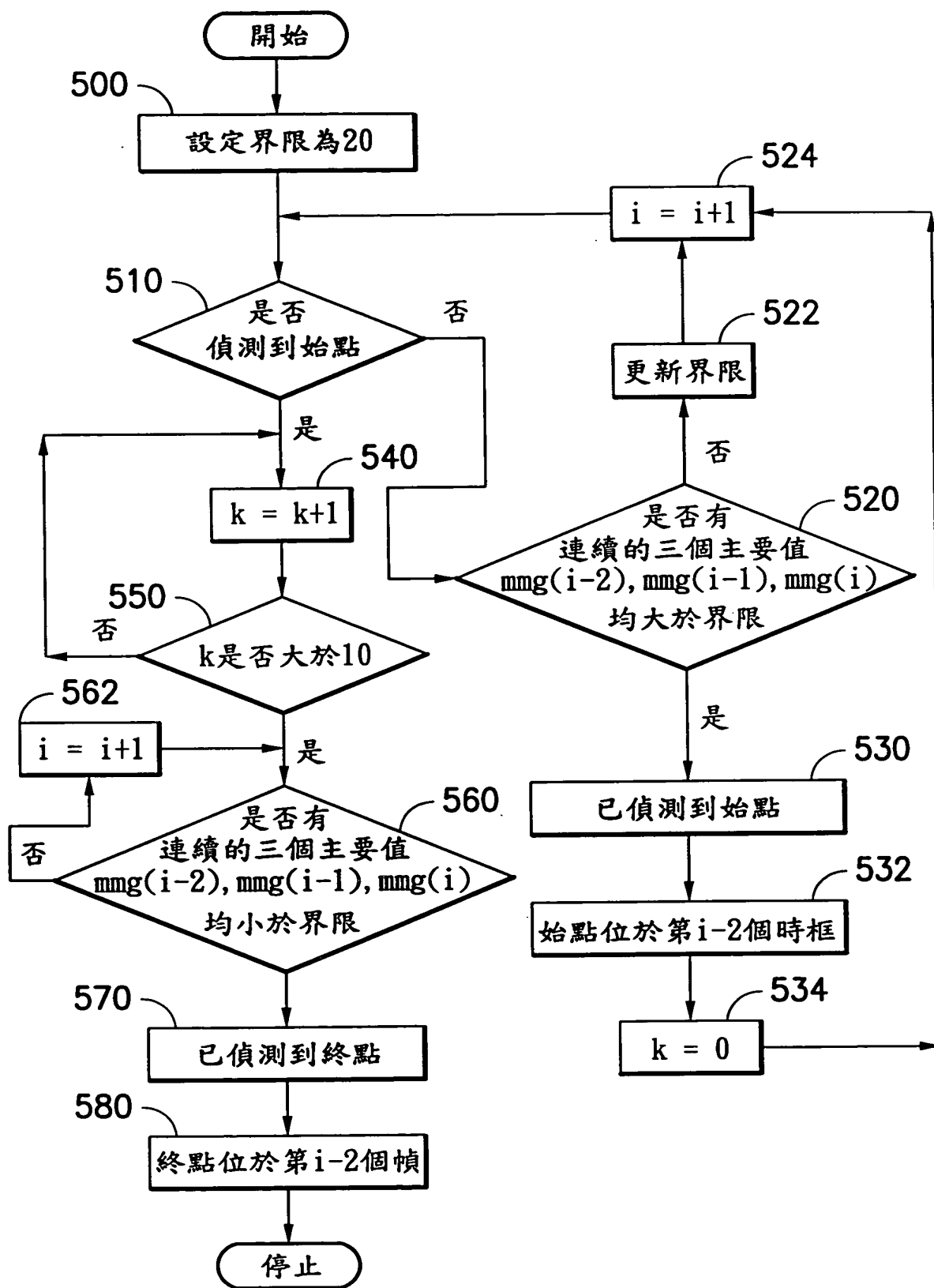
第 2 圖



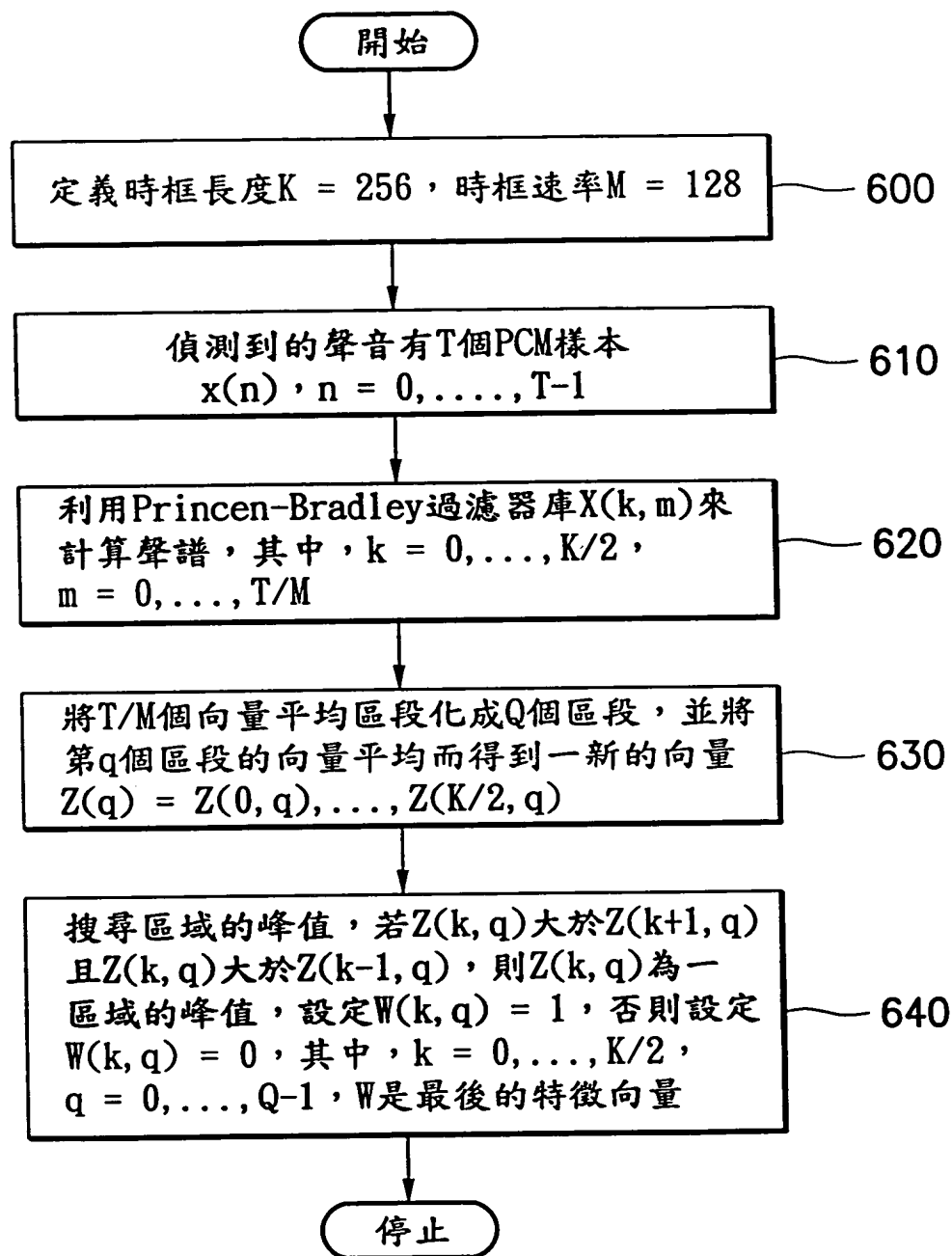
第 3 圖



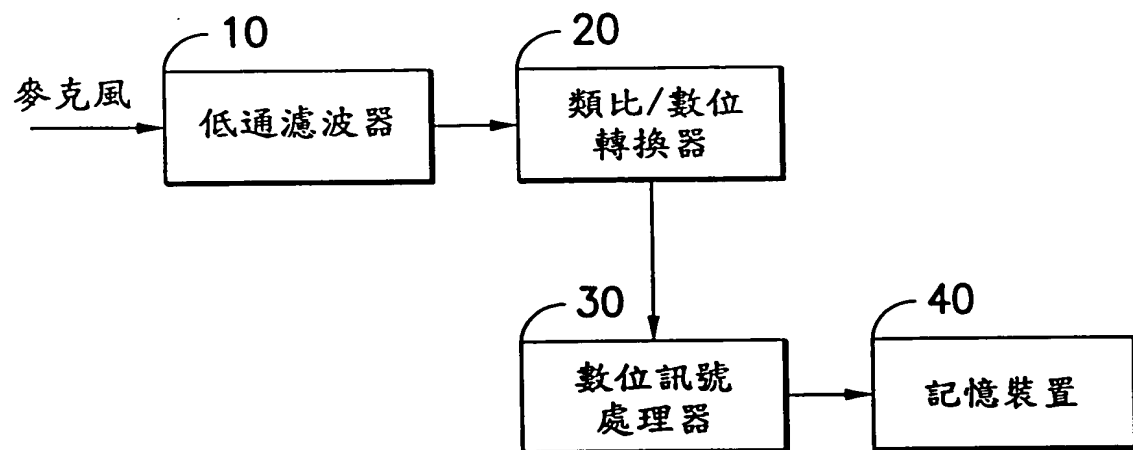
第 4 圖



第 5 圖

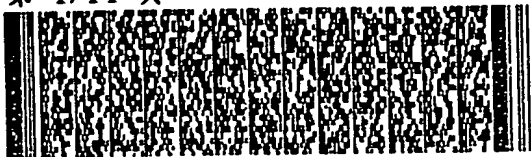


第 6 圖

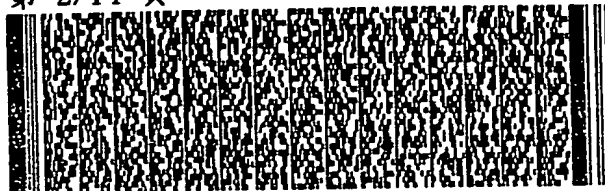


第 7 圖

第 1/14 頁



第 2/14 頁



第 4/14 頁



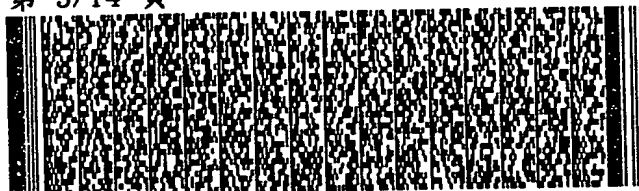
第 4/14 頁



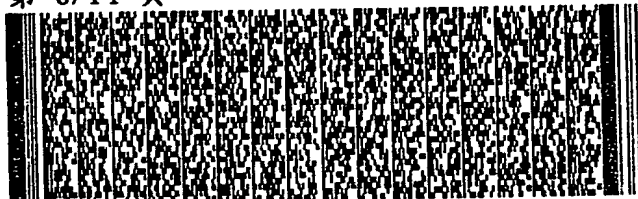
第 5/14 頁



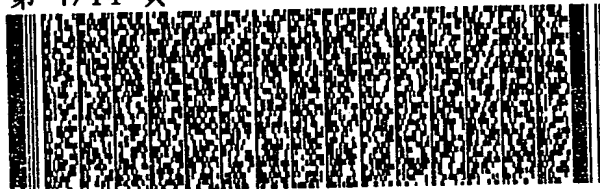
第 5/14 頁



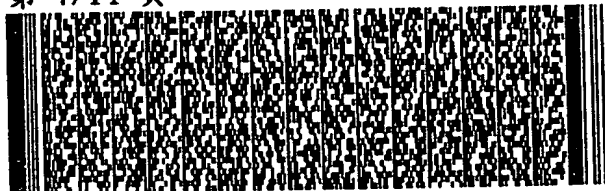
第 6/14 頁



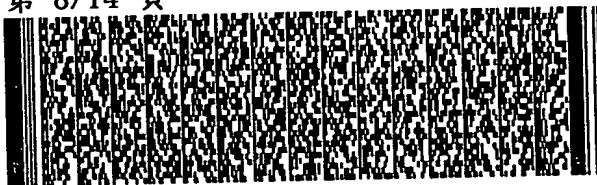
第 7/14 頁



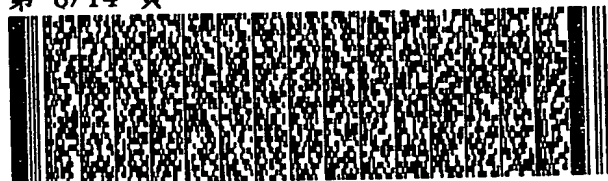
第 7/14 頁



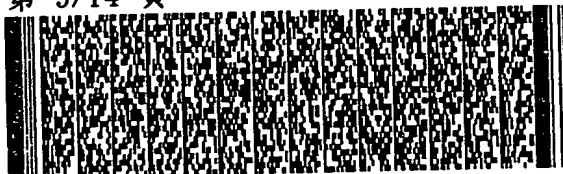
第 8/14 頁



第 8/14 頁



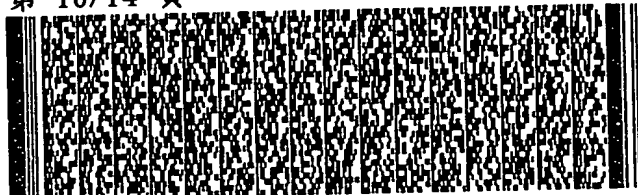
第 9/14 頁



第 9/14 頁



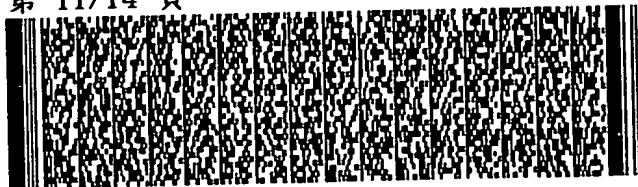
第 10/14 頁



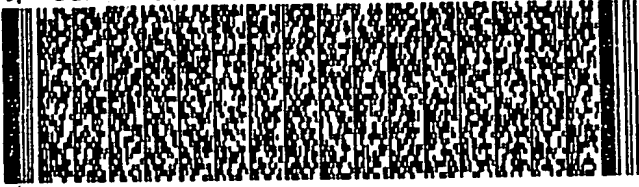
第 10/14 頁



第 11/14 頁



第 11/14 頁



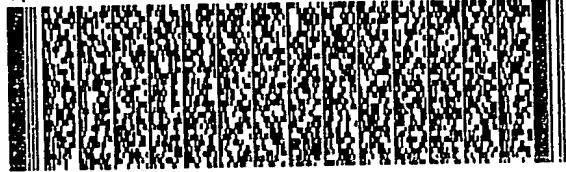
第 12/14 頁



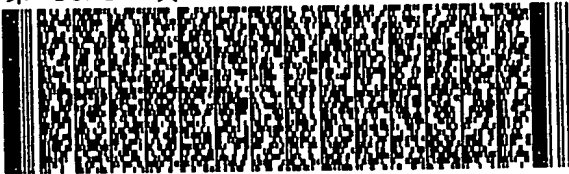
第 12/14 頁



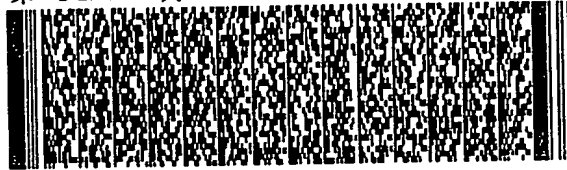
第 13/14 頁



第 13/14 頁



第 14/14 頁



第 14/14 頁

